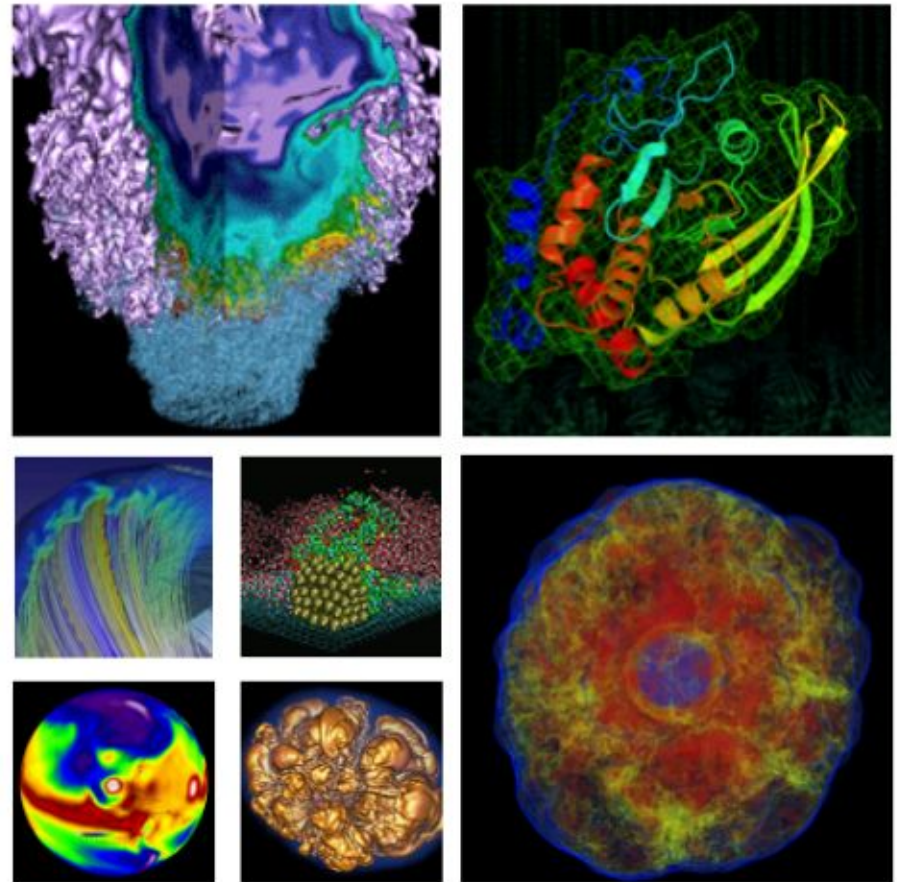


Doing Machine Learning @ NERSC



Evan Racah
New User Training

2/24/17

What is Machine Learning?



Formal Definition: “A form of applied statistics with increased emphasis on the use of computers to statistically estimate complicated functions” (Goodfellow. et al.)

What is Machine Learning?



Basically instead of specifying an exact function, we use data to learn the best function for the task at hand

Example: Linear Regression!

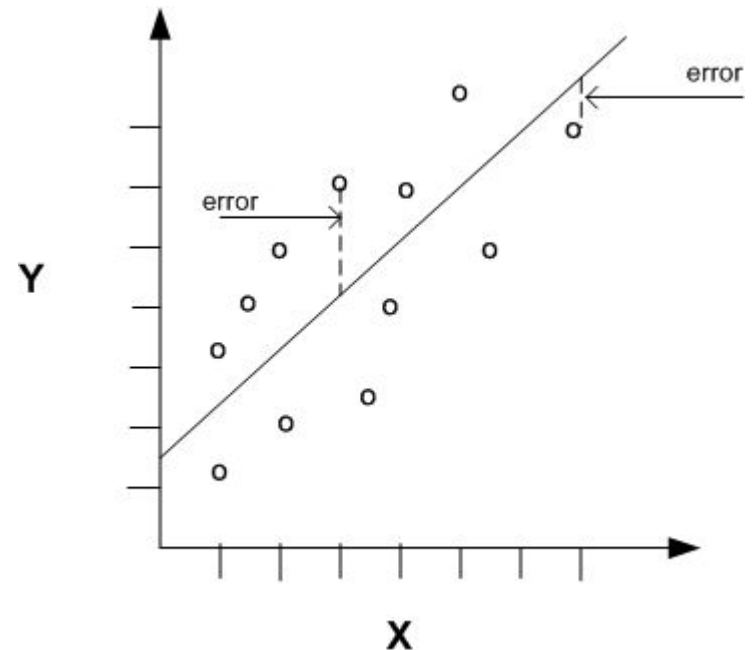


Data

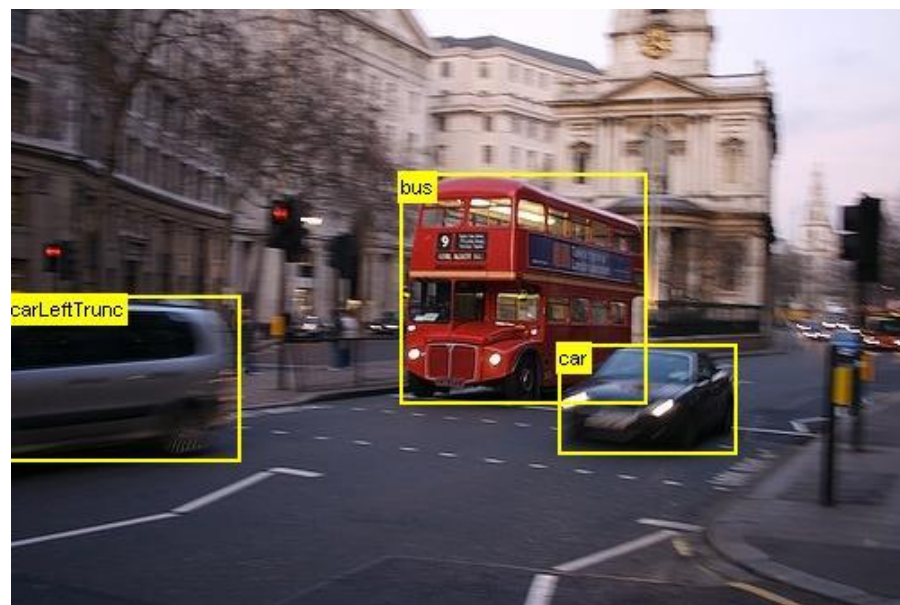
- Set of x, y pairs

Function

- The optimal linear function ($y = \mathbf{W}x + \mathbf{b}$) that best fits the pairs



More Advanced Examples/Applications



Why Might You Need Machine Learning?

- You are trying to predict or assign a value to each instance of your data
- This can't be done easily or efficiently programmatically
- A human can't do it or it would be too tedious for them to
- You have many existing examples of pairs of a data instance and the corresponding true value

Tools



Python library: one-stop shop for doing machine learning

Easy to use, great coverage of ML algorithms



Automatic parallelization of your code.

Some support for ML algorithms.



General purpose library for numerical computation, but mostly used for deep learning.

Large community, very popular



Keras

Deep learning library built on top of Theano or TensorFlow
Can design and train neural net in fewer than 20 lines of code.

The Caffe logo is located in the lower middle right section. It features the word "Caffe" in a red, serif font, set against a light beige rectangular background.

Built with computer vision in mind

Very off the shelf use cases require no programming (just specifying a config file)

How Do I Use These Tools at NERSC?

Deep Learning Module

- Python deep learning tools available under the deep learning module
- Just one module load call and then they are ready to be imported in your python script!

```
racah@cori04:~> module load deeplearning
racah@cori04:~> █
```

Scikit-Learn

- available in standard python module and deep learning one

```
racah@cori03:~> module load caffe
```

Caffe and Spark

- available as separate modules

```
racah@cori03:~> module load spark
```

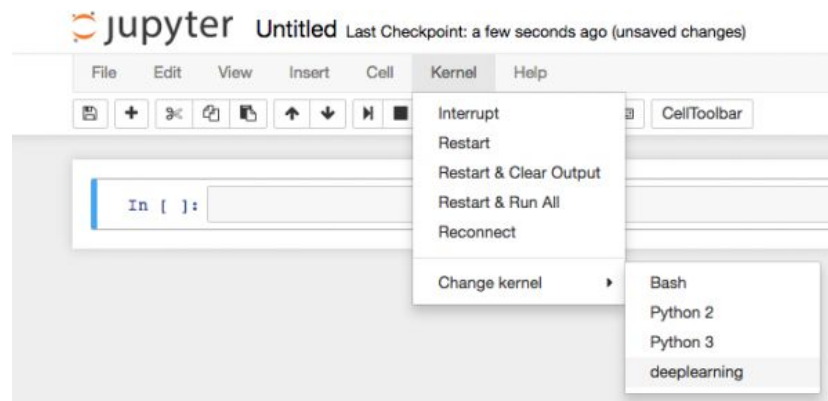
Interactive computing:

Q: You a big Jupyter notebook fan?!

A: No problem. The iPython deeplearning kernel allows for interactively using the deep learning module python tools

For more information visit

<http://www.nersc.gov/users/data-analytics/data-analytics/deep-learning/>



Demo

<https://github.com/NERSC/new-user-training-notebooks>



NERSC

National Energy Research Scientific
Computing Center